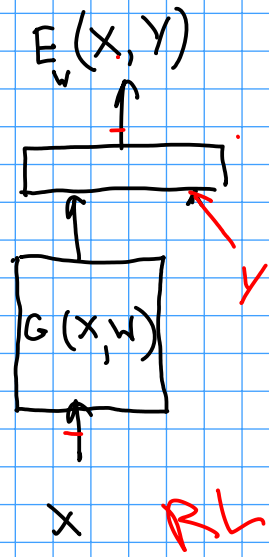
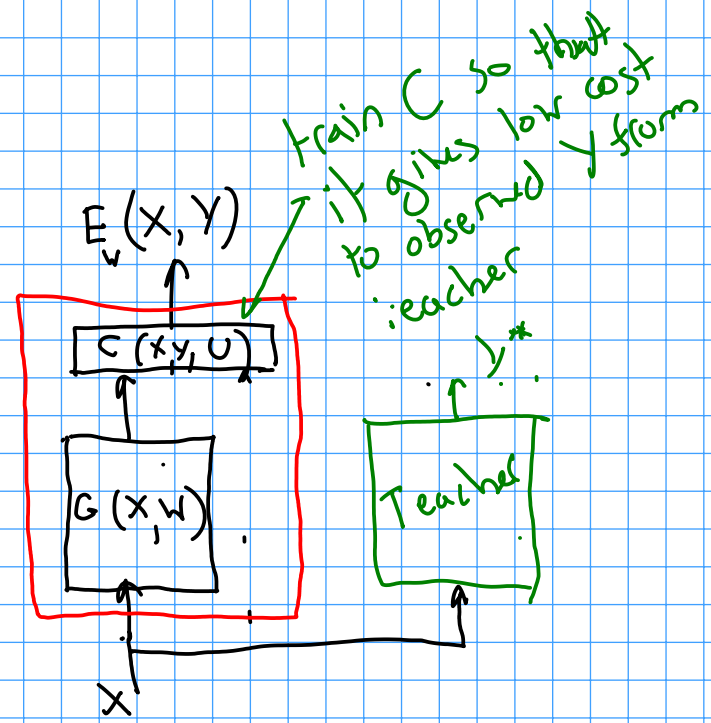
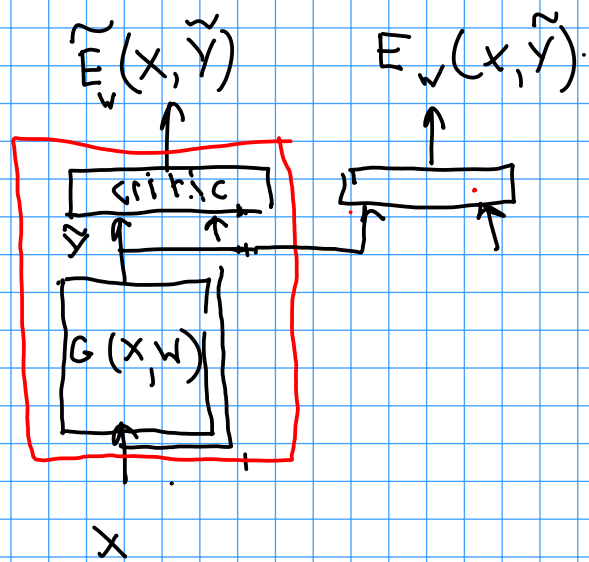


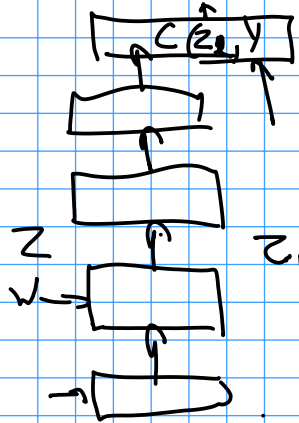
Unknown  
Not differentiable  
Black-box  
optimization  
REINFORCE  
Williams



Actor-critic.



Imitation learning  
Inverse RL



$x^i, y^i$

$$z_k = F_k(z_{k-1}, w_k)$$

$$L(z^i, y^i, w) = c(z_l^i, y^i) + \sum_k \lambda_k^T (z_k - F_k(z_{k-1}, w_k))$$

$$\mathcal{L}(w) = \sum_i L(z^i, y^i, w)$$

$$\rightarrow \frac{\partial \mathcal{L}}{\partial \lambda_k^T} = 0 \Rightarrow z_k = F_k(z_{k-1}, w_k) \quad \text{prop}$$

$$\begin{aligned} \rightarrow \frac{\partial \mathcal{L}}{\partial z_k} &= 0 \Rightarrow \lambda_k = \frac{\partial F_{k+1}}{\partial z_k} \lambda_{k+1} \\ \rightarrow \frac{\partial \mathcal{L}}{\partial w_k} &= 0 \end{aligned} \quad \text{adjoint system}$$